

# A CENTRAL LIMIT THEOREM FOR FLUCTUATIONS IN POLYANALYTIC GINIBRE ENSEMBLES

ANTTI HAIMI AND ARON WENNMAN

**ABSTRACT.** We study fluctuations of linear statistics in Polyanalytic Ginibre ensembles, a family of point processes describing planar free fermions in a uniform magnetic field at higher Landau levels. Our main result is asymptotic normality of fluctuations, extending a result of Rider and Virág. As in the analytic case, the variance is composed of independent terms from the bulk and the boundary. Our methods rely on a structural formula for polyanalytic polynomial Bergman kernels which separates out the different pure  $q$ -analytic kernels corresponding to different Landau levels. The fluctuations with respect to these pure  $q$ -analytic Ginibre ensembles are also studied, and a central limit theorem is proved. The results suggest a stabilizing effect on the variance when the different Landau levels are combined together.

## 1. INTRODUCTION

The Ginibre ensemble is one of the major point processes in random matrix theory and mathematical physics. The model has at least three possible interpretations: in terms of eigenvalues of random matrices, Coulomb gas of charged particles in an external field, or ground state free fermions in a magnetic field perpendicular to the plane. In this paper, we study *Polyanalytic Ginibre ensembles* [12], a family of point process which generalizes the last of these three notions so that the particles are allowed to occupy more general energy levels.

From the point of view of quantum mechanics, we can arrive at the Polyanalytic Ginibre ensembles in the following way. It is a consequence of the Pauli exclusion principle that the wavefunction of the  $N$ -body system of free (i.e. non-interacting) particles is given by the Slater determinant

$$(1.1) \quad \det[\psi_j(z_j)]_{1 \leq i, j \leq N}$$

where  $\psi_j$  are orthonormal single particle wave functions. We take these to be eigenstates of the *Landau Hamiltonian*

$$H_B := \frac{1}{2} \left( \left( \frac{\partial}{\partial x} - \frac{B}{2}y \right)^2 + \left( i \frac{\partial}{\partial y} - \frac{B}{2}x \right)^2 \right),$$

---

A.H. was supported by Lise Meitner grant of Austrian Science Fund (FWF).

which is known to describe a single electron in the plane subjected to a perpendicular magnetic field of strength  $B$ . It is known ([10], [1]) that the spectrum of this operator (acting on  $L^2(\mathbb{R}^2)$ ) consists of eigenvalues (so called *Landau levels*)

$$e_q^B = (q + 1/2)B$$

and the eigenspace corresponding to  $e_q^B$  consists of functions of the form

$$f(z)e^{-B|z|^2/2},$$

where  $f$  belongs to *pure  $q$ -analytic Bargmann-Fock space*  $A_{\delta;B,q}^2$ , defined as the orthogonal difference  $A_{B,q}^2 \ominus A_{B,q-1}^2$  between two consecutive  *$q$ -analytic Bargmann-Fock spaces*

$$A_{B,q}^2 := \left\{ f : \int_{\mathbb{C}} |f(z)|^2 e^{-B|z|^2} dA(z) < \infty, \bar{\partial}^q f(z) = 0 \right\}.$$

The case of the Ginibre ensemble (i.e.  $q = 1$ ) corresponds to the classical Bargmann-Fock space, which is associated with the orthonormal basis

$$\psi_j(z) = \frac{B^{1/2}}{\sqrt{j!}} (B^{1/2}z)^j e^{-B|z|^2/2}, \quad 0 \leq j.$$

According to a well-known computation in point process theory, the probability density corresponding to the many-body wavefunction in (1.1) can be written (up to a constant) in a determinantal form:

$$\left| \det[\psi_j(z_j)]_{1 \leq i,j \leq N} \right|^2 \sim \frac{1}{N!} \det[K(z_i, z_j)]_{1 \leq i,j \leq N} e^{-B|z_1|^2 - \dots - B|z_N|^2},$$

where

$$K(z, w) = \sum_{j=0}^{N-1} \varphi_j(z) \overline{\varphi_j(w)}.$$

In the Ginibre case, this coincides with the reproducing kernel of the space of analytic polynomials of degree  $\leq N - 1$  in  $L^2(e^{-B|z|^2} dA(z))$ .

To obtain the Polyanalytic Ginibre ensembles, we allow wavefunctions from general eigenspaces of  $H_B$ , not just the from lowest level. The point processes that we will consider fall into two categories: *full type* and *pure type*. In the first case, we have  $n$  particles at each level up to  $q$  and in the second  $n$  particles at  $q$ 'th level only. So, the processes of full type contain  $nq$  points, and those of the pure type consist of  $n$  points. It is natural to take the field  $B$  to be equal to the number of particles  $n$  at each level; physically, this corresponds to each level being completely filled. The corresponding reproducing kernels are formed by choosing the appropriate wavefunctions from  $A_{B,q}^2$  and  $A_{\delta;B,q}^2$  and will be denoted by  $K_{n,q}$  and  $K_{\delta;n,q}$ . They correspond to spaces

$$\text{Pol}_{n,q} = \text{span}\{\bar{z}^r z^j : 0 \leq j \leq n - 1, 0 \leq r \leq q - 1\} \subset L^2(e^{-n|z|^2} dA)$$

and

$$\delta\text{Pol}_{n,q} := \text{Pol}_{n,q} \ominus \text{Pol}_{n,q-1},$$

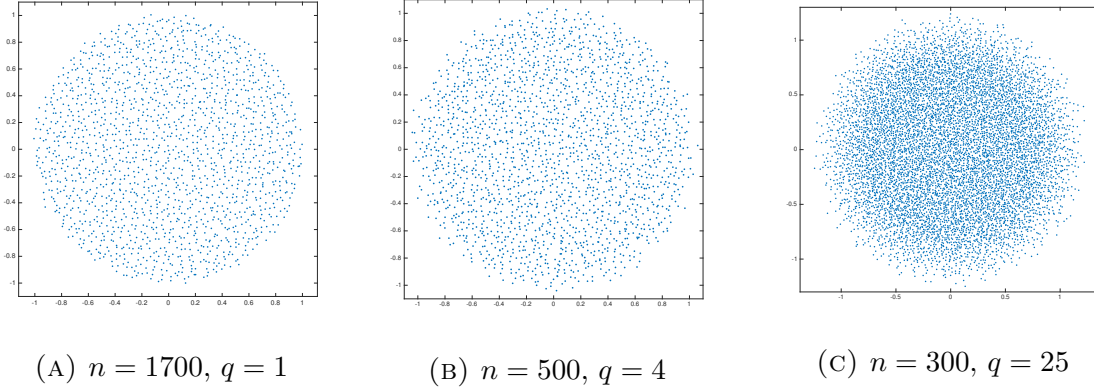


FIGURE 1. Polyanalytic Ginibre ensembles defined by the kernel  $K_{n,q}$  with different values of  $q$  and  $n$ . Notice that when  $q$  is relatively large compared to  $n$ , the density is less uniform.

respectively. We see that when we allow higher Landau levels in the process, the function spaces do not only consist of analytic functions as in the Ginibre case, but rather more general *polyanalytic functions*. The study of these functions has attracted increasing attention recently, and interesting connections between signal analysis, quantum mechanics and complex analysis have been found. For an introduction, see [3] or [1].

The study of processes of this type was initiated in mathematics literature in [12], where precise estimates and scaling limits for the kernels  $K_{\delta;n,q}$  and  $K_{n,q}$  were obtained when  $n \rightarrow \infty$  and  $q$  is fixed. In particular, it follows from the results there that the circular law will appear in the limit also when higher Landau levels are included (see Figure 1). Our main theorem shows asymptotic normality of fluctuations around this mean, generalising a result of Rider and Virág [15] from the analytic case. For a continuous test function, we define the linear statistics associated with processes of the pure type by

$$\text{trace}_{\delta;n,q} f = \sum_{j=1}^n f(\lambda_j),$$

where the vector  $(\lambda_j)_{j=1}^n$  is picked from the determinantal process defined by the kernel  $K_{\delta;n,q}$ . The random variables  $\text{trace}_{n,q}$  are defined similarly:

$$\text{trace}_{n,q} f = \sum_{j=1}^{nq} f(\lambda_j).$$

Here we have taken  $nq$  points from the process defined by  $K_{n,q}$ . The corresponding fluctuations are defined by

$$\text{fluct}_{\delta;n,q} = \text{trace}_{\delta;n,q} - \mathbb{E}(\text{trace}_{\delta;n,q}), \quad \text{fluct}_{n,q} = \text{trace}_{n,q} - \mathbb{E}(\text{trace}_{n,q}).$$

We denote by  $N(a, \sigma^2)$  a normal variable with mean  $a$  and variance  $\sigma^2$ .

**Theorem 1.1.** *Let  $g \in C_0^\infty(\mathbb{C})$  be real-valued. We have*

$$(1.2) \quad \text{fluct}_{\delta;q,n} g \rightarrow N\left(0, (2q-1)\|g\|_{H^1(\mathbb{D})}^2 + \frac{1}{2}\|g\|_{H^{1/2}(\partial\mathbb{D})}^2\right),$$

and

$$(1.3) \quad \text{fluct}_{q,n} g \rightarrow N\left(0, q(\|g\|_{H^1(\mathbb{D})}^2 + \frac{1}{2}\|g\|_{H^{1/2}(\partial\mathbb{D})}^2)\right),$$

in distribution as  $n \rightarrow \infty$ . Here,

$$\|g\|_{H^1(\mathbb{D})}^2 := \int_{\mathbb{D}} |\bar{\partial}g|^2 dA(z)$$

is the Dirichlet semi-norm of  $g$  on  $\mathbb{D}$  and

$$\|g\|_{H^{1/2}(\partial\mathbb{D})}^2 = \sum_{k \in \mathbb{Z}} |k| |\hat{g}(k)|^2,$$

is the  $H^{1/2}$  semi-norm of  $g$  on the unit circle. The numbers

$$\hat{g}(k) := \frac{1}{2\pi} \int_0^{2\pi} g(e^{i\theta}) e^{-ik\theta} d\theta$$

are the Fourier coefficients of the  $g$  restricted to  $\{|z| = 1\}$ .

We observe that the variance consists of essentially the same two terms as in the case  $q = 1$ , the bulk term and the boundary term. However, the coefficients in front of the two terms bring in new phenomena. In the formula (1.3), the variance from the analytic case is just multiplied by  $q$ . In the pure polyanalytic case (1.2), only the bulk term involves a factor depending on  $q$  and the boundary is the same as in the analytic setting. Moreover, if we concentrate only on test functions which are supported in the bulk, the variance in (1.2) is higher than in (1.3). In fact, because  $\frac{1}{q} \sum_{r=1}^q (2r-1) = q$ , the variance in (1.3) is obtained by averaging the variances from each of the  $q$  Landau levels. This suggests that adding several Landau levels together has a certain smoothening effect on the variance. Interestingly, for the boundary terms the situation is different, because the boundary contribution in (1.3) is just the sum of boundary contributions from the individual levels. We do not have a physical explanation for these facts at the moment.

As in [15] and [6], the proof is based on the cumulant method introduced by Costin and Lebowitz [9]. Otherwise the argument is quite different. Instead of using explicit expression for the correlation kernel or estimates of it directly, our proof starts with a partial integration procedure (Proposition 3.1), based on expressing the kernels in terms of quantum mechanical raising operators which act isometrically between Landau level eigenspaces. As a result, we can rewrite the cumulants in terms of the Ginibre kernel only. This representation combined with rather general estimates of integrals involving cyclic products of kernels allows us to make a reduction to the analytic situation in [15]. We want to emphasize that even though precise

estimates and an explicit expressions for our polyanalytic kernels are known, applying formulas for the cumulants directly would most likely lead to be extremely complicated calculations.

In this paper, the function spaces are defined with a Gaussian weight only. In the analytic case, the result of Rider and Virág has been generalised to more general weights by Ameur, Hedenmalm and Makarov (in [6] for test functions with support in the bulk, and in [5] for general test functions). While it is reasonable to expect that results of the former type can be extended to polyanalytic setting by using estimates from [11], it remains an interesting question what happens with general test functions. The technique in [5] seems not to generalize to our polyanalytic case. On the other hand, our approach is based on expressing the polyanalytic kernels in terms of the Landau level raising operators and these are closely tied to the Gaussian case.

Our methods could be used to study other configurations of particles as well, not just those where either one level or all levels up to a given level are filled. However, our techniques do require the highest level  $q$  to be fixed as we let  $n \rightarrow \infty$ . In a forthcoming paper we will study asymptotic behaviour of the kernel  $K_{n,q}$  when both  $q$  and  $n$  tend to infinity. Some preliminary observations about this setting can be found in [12].

The paper is organised as follows. In Section 2, we provide basic facts about the function spaces and point processes involved. In Section 3, we introduce cumulants and our main technique, the partial integration procedure in Proposition 3.1. In Section 4, we prove certain estimates for integrals involving cyclic products of kernels that allow us to estimate the cumulants. The main theorem is then proved in Section 5.

**Acknowledgements.** We would like to thank Joaquim Ortega-Cerdà for sharing the code which was used to produce the simulations in Figure 1 and Seung-Yeop Lee for interesting remarks concerning the main theorem.

## 2. PRELIMINARIES

**2.1. Notation.** We write  $z = x + iy$  and let  $\partial = \frac{1}{2}(\partial_x - i\partial_y)$  and  $\bar{\partial} = \frac{1}{2}(\partial_x + i\partial_y)$  denote the standard Wirtinger differential operators. We will write  $\mathbb{D}$  for the open unit disk. We will write  $dA(z) = \frac{1}{\pi}dx dy$  and let  $d\mu_n(z) = e^{-n|z|^2}dA(z)$  be the Gaussian measure on  $\mathbb{C}$ . We will use the notation  $d\mu_n^k(z_1, \dots, z_k) = d\mu_n(z_1) \dots d\mu_n(z_k)$ .

**2.2. Spaces of polyanalytic polynomials.** A function  $f$  defined on a subset of the complex plane is called  $q$ -analytic if it satisfies the equation  $\bar{\partial}^q f = 0$  in the sense of distributions.

Equivalently, a function  $f$  is  $q$ -analytic if it can be decomposed as

$$f(z) = \sum_{j=0}^{q-1} \bar{z}^j f_j(z),$$

where the functions  $f_j$  are analytic. A function that is  $q$ -analytic for some  $q$  is called *polyanalytic*. We define the *Bargmann-Fock space of  $q$ -analytic functions* as

$$A_{n,q}^2 := \left\{ f(z) : \int_{\mathbb{C}} |f(z)|^2 d\mu_n(z), \bar{\partial}^q f = 0 \right\}.$$

The spaces of  *$q$ -analytic polynomials* are defined by

$$\text{Pol}_{n,q} = \text{span} \left\{ \bar{z}^r z^j : 0 \leq j \leq n-1, 0 \leq r \leq q-1 \right\}$$

We equip  $\text{Pol}_{n,q}$  with the inner product from  $A^2(e^{-n|z|^2})$ . The *Bargmann-Fock space of pure  $q$ -analytic functions* is defined as the orthogonal difference

$$A_{\delta;n,q}^2 := A_{n,q}^2 \ominus A_{n,q-1}^2.$$

The spaces of *pure  $q$ -analytic polynomials* are defined analogously by

$$\delta\text{Pol}_{n,q} = \text{Pol}_{n,q} \ominus \text{Pol}_{n,q-1}.$$

Any function space  $H$  encountered here possesses a reproducing kernel, i.e. a function  $K(z, w)$  such that for any  $w \in \mathbb{C}$ ,  $K_w := K(\cdot, w)$  is an element of the space, and for any function  $f \in H$  it holds that

$$f(w) = \langle f, K_w \rangle, \quad w \in \mathbb{C}.$$

We denote by  $K_{n,q}(z, w)$  the reproducing kernel for the space  $\text{Pol}_{n,q}$  and by  $K_{\delta;n,q}(z, w)$  the kernel for the space  $\delta\text{Pol}_{n,q}$ . The kernel  $K_{n,1}(z, w)$  is simply written  $K_n(z, w)$ . Clearly,  $K_{n,q}$  can be expressed in terms of the pure  $q$ -analytic kernels as follows:

$$(2.1) \quad K_{n,q}(z, w) = \sum_{r=1}^q K_{\delta;n,r}(z, w).$$

In [12], the following explicit expression for the kernel  $K_{n,q}$  was given:

$$(2.2) \quad \begin{aligned} K_{n,q}(z, w) = & m \sum_{r=0}^{q-1} \sum_{i=0}^{n-r-1} \frac{r!}{(r+i)!} (nz\bar{w})^i L_r^i(n|z|^2) L_r^i(n|w|^2) \\ & + n \sum_{j=0}^{q-2} \sum_{k=1}^{q-j-1} \frac{j!}{(k+j)!} (n\bar{z}w)^k L_j^k(n|z|^2) L_j^k(n|w|^2). \end{aligned}$$

The kernel  $K_{\delta;n,q}$  can then be obtained from this and the relation

$$(2.3) \quad K_{\delta;n,q} = K_{n,q}(z, w) - K_{n,q-1}(z, w).$$

We will not need to use the explicit expressions for these kernels in this paper. We just observe that

$$K_n(z, w) = K_{n,1}(z, w) = n \sum_{j=0}^{n-1} \frac{(nz\bar{w})^j}{j!}$$

is the standard Ginibre kernel from random matrix theory. Rather than use (2.2) and (2.3), we will express  $K_{n,q}$  and  $K_{\delta;n,q}$  in terms of  $K_n$  and the *raising operators*

$$(T_{n,r}f)(z) = \frac{n^{-r/2}}{\sqrt{r!}} e^{n|z|^2} \partial^r \left\{ f(z) e^{-n|z|^2} \right\}.$$

Frequently, it will be convenient to use the notation  $T_n := -n^{1/2}T_{n,1}$ . As explained in [12],  $T_{n,r}$  is an isometric isomorphism from  $\text{Pol}_{n,1}$  to  $\delta\text{Pol}_{n,r+1}$ . Thus,  $T_{n,r}$  maps orthonormal bases to orthonormal bases. It thus implies that the pure polyanalytic kernels may be obtained from the analytic kernel  $K_n(z, w)$  by

$$K_{\delta;n,r}(z, w) = [T_{n,r-1}]_z \overline{[T_{n,r-1}]_w} K_n(z, w).$$

Consequently,

$$(2.4) \quad K_{n,q}(z, w) = \sum_{r=0}^{q-1} [T_{n,r}]_z \overline{[T_{n,r}]_w} K_n(z, w) = \sum_{r=0}^{q-1} \frac{n^{-r}}{r!} [T_n]_z^r \overline{[T_n]_w^r} K_n(z, w).$$

Polyanalytic functions appear naturally in e.g. quantum mechanics and time–frequency analysis. The reader who is interested in more background can consult [3], [2], [18] or [8].

**2.3. Polyanalytic Ginibre ensembles.** Our aim is to study determinantal processes associated with the kernels  $K_{n,q}$  and  $K_{\delta;n,q}$ . These are given by the following probability measures on  $\mathbb{C}^{nq}$  and  $\mathbb{C}^n$ :

$$d\mathbb{P}_{n,q} := \frac{1}{(nq)!} \det[K_{q,n}(z_j, z_k)]_{1 \leq j,k \leq nq} e^{-n|z_1|^2 - \dots - n|z_{nq}|^2} dA(z_1) \dots dA(z_{nq})$$

and

$$d\mathbb{P}_{\delta;n,q} := \frac{1}{n!} \det[K_{\delta;q,n}(z_j, z_k)]_{1 \leq j,k \leq n} e^{-n|z_1|^2 - \dots - n|z_n|^2} dA(z_1) \dots dA(z_n).$$

The fact that these are probability measures is a standard result in theory of determinantal point processes (see e.g. [13] for an introduction). We will identify all copies of  $\mathbb{C}$  and interpret  $d\mathbb{P}_{n,q}$  and  $d\mathbb{P}_{\delta;n,q}$  as densities for random configurations of  $nq$  and  $n$  unlabelled points in  $\mathbb{C}$ , respectively. It is well known ([14], [17]) that any locally trace class projection kernel defines a determinantal process. Because our spaces are finite dimensional, we do not need this general definition here. We just note that an infinite dimensional counterpart of polyanalytic Ginibre ensembles has been studied by Shirai [16] and it belongs to a more general class of point processes called *Weyl–Heisenberg ensembles*, recently introduced in [4].

**2.4. Linear statistics.** Recall from the introduction that given a bounded, continuous function  $g$ , the linear statistics are defined by

$$\text{trace}_{n,q}g := \sum_{j=1}^{nq} g(\lambda_j),$$

where  $(\lambda_1, \dots, \lambda_{nq})$  is picked from the measure  $d\mathbb{P}_{n,q}$ . The variables  $\text{trace}_{\delta;n,q}g$  are defined in a similar way, but one picks a random vector  $(\lambda_1, \dots, \lambda_n)$  from  $d\mathbb{P}_{\delta;n,q}$  instead.

Asymptotic behaviour of the expectation of linear statistics can be analysed as follows:

$$\frac{1}{nq} \mathbb{E} \text{trace}_{q,n}(f) = \mathbb{E} f(z_1) = \frac{1}{nq} \int_{\mathbb{C}} f(z) K_{q,n}(z, z) e^{-n|z|^2} dA(z) \rightarrow \int_{\mathbb{D}} f(z) dA.$$

Here, the second equality is a general fact about determinantal point processes. The limit follows from the weak convergence  $\frac{1}{nq} K_{q,n}(z, z) e^{-n|z|^2} \rightarrow 1_{\mathbb{D}}$  which follows from the results in [12]. A similar statement is true for pure polyanalytic kernels  $K_{\delta;n,q}$ . This generalizes the well-known circular law about the Ginibre ensemble (the case  $q = 1$ ). The interpretation is that the particles tend to accumulate uniformly on the unit disk. We use standard terminology and refer to the open unit disk as the *bulk*.

The goal of this paper is to understand how the linear statistics fluctuate around the mean. We let  $\text{fluct}_{n,q}$  and  $\text{fluct}_{\delta;n,q}$  be the mean-zero variables

$$\text{fluct}_{n,q}g = \sum_{\lambda_j} g(\lambda_j) - \mathbb{E}g(\lambda_j).$$

where  $\lambda_1, \dots, \lambda_{nq}$  is picked from the density  $d\mathbb{P}_{n,q}$ . The variables  $\text{fluct}_{\delta;n,q}g$  have an analogous definition.

### 3. CUMULANTS

For any real-valued random variable  $X$ , the cumulants  $C_k(X)$  are defined implicitly by

$$\log \mathbb{E} [e^{tX}] = \sum_{j=1}^{\infty} \frac{t^j}{j!} C_j(X).$$

The first cumulant is equal to the expectation and the second to the variance of  $X$ . For linear statistics of determinantal processes, the cumulants are given by an explicit formula introduced by Costin and Lebowitz in [9]. We will use a formulation from [6]. For  $X = \text{trace}_{n,q}(g)$ , we have

$$(3.1) \quad C_k(X) = \int_{\mathbb{C}^k} G_k(z_1, \dots, z_k) K_{n,q}(z_1, z_2) K_{n,q}(z_2, z_3) \cdots K_{n,q}(z_k, z_1) d\mu_n^k(z_1, \dots, z_k),$$



where

$$G_k(z_1, \dots, z_k) = \sum_{j=1}^k \frac{(-1)^{j-1}}{j} \sum_{\substack{k_1+k_2+\dots+k_j=k \\ k_1, \dots, k_j \geq 1}} \frac{k!}{k_1!k_2!\dots k_j!} \prod_{l=1}^j g(z_l)^{k_l}.$$

For  $X = \text{trace}_{\delta;n,q}(g)$ , one just replaces each occurrence of  $K_{n,q}$  by  $K_{\delta;n,q}$  in (3.1). For future reference, note that  $G_k$  is a sum of products of functions of one variable.

It is known that convergence of the cumulants implies convergence in distribution. Moreover, a random variable is normally distributed iff all cumulants of order  $k \geq 3$  vanish. Therefore, to prove the main theorem, we need to show that for  $k \geq 3$ ,  $C_k(\text{trace}_{n,q})$  and  $C_k(\text{trace}_{\delta;n,q})$  tend to zero and compute the limits of the second cumulants as  $n \rightarrow \infty$ .

The following proposition gives an alternative way to represent the cumulants. In this new representation, the cumulants are written as integrals involving a cyclic product with the analytic Ginibre kernel  $K_n$ .

For non-negative integers  $\alpha, \beta$  and  $n$ , we define the differential operators

$$\mathcal{D}_{\alpha,\beta,n} = \sum_{j=0}^{\min(\alpha,\beta)} \binom{\min(\alpha,\beta)}{j} (\max(\alpha,\beta))_j n^j \bar{\partial}^{\alpha-j} \partial^{\beta-j},$$

where  $(r)_j = r(r-1)\dots(r-j+1)$ . Assuming  $\beta \geq \alpha$ , this can be written in a slightly more compact form as

$$\mathcal{D}_{\alpha,\beta,n} = \alpha! n^\alpha \bar{\partial}^{\beta-\alpha} L_\alpha^{\beta-\alpha}(-n^{-1}\Delta),$$

where

$$L_r^k := \sum_{j=0}^r (-1)^j \binom{r+k}{r-j} \frac{1}{j!} x^j$$

denotes the associated Laguerre polynomial of index  $k$  and degree  $r$ . Similarly, when  $\alpha \geq \beta$ , we have

$$\mathcal{D}_{\alpha,\beta,n} = \beta! n^\beta \bar{\partial}^{\alpha-\beta} L_\beta^{\alpha-\beta}(-n^{-1}\Delta),$$

In particular, in the special case  $\alpha = \beta = r$  we have

$$(3.2) \quad \mathcal{D}_{\alpha,\beta,n} = n^r r! L_r^0(-n^{-1}\Delta).$$

**Proposition 3.1.** *Let  $F : \mathbb{C}^k \rightarrow \mathbb{C}$  be bounded and smooth. We have*

$$\begin{aligned} & \int_{\mathbb{C}^k} F(z_1, \dots, z_k) [T_n]_{z_1}^{i_1} \overline{[T_n]_{z_2}^{i_1}} K(z_1, z_2) \dots [T_n]_{z_k}^{i_k} \overline{[T_n]_{z_1}^{i_k}} K(z_k, z_1) d\mu^k(z_1, \dots, z_k) \\ &= \int_{\mathbb{C}^k} \mathcal{D}_{i_1, \dots, i_k, n} F(z_1, \dots, z_k) K_n(z_1, z_2) \dots K_n(z_k, z_1) d\mu^k(z_1, \dots, z_k). \end{aligned}$$

where

$$\mathcal{D}_{i_1, \dots, i_k, n} = [\mathcal{D}_{i_k, i_1, n}]_{z_1} [\mathcal{D}_{i_1, i_2, n}]_{z_2} \dots [\mathcal{D}_{i_{k-1}, i_k, n}]_{z_k}.$$

*Proof.* We will use that

$$(3.3) \quad \int f(z) T_n g(z) d\mu_n(z) = \int \partial f(z) g(z) d\mu_n(z)$$

where  $f, g \in C^1$ . In our applications,  $f$  and  $g$  will have at most polynomial growth in  $z$ , so boundary terms do not appear in the partial integration. We will also need the following: for  $j \leq r$ ,

$$(3.4) \quad \bar{\partial}^j T_n^r f = (r)_j n^j T_n^{r-j} f$$

for an analytic function  $f$ . This can be seen as follows:

$$\begin{aligned} \bar{\partial}^j T_n^r f &= (-1)^r \bar{\partial}^j \sum_{k=0}^r \binom{r}{k} f^{(k)}(z) (-n\bar{z})^{r-k} \\ &= (-1)^{r+j} n^j \sum_{k=0}^{r-j} \binom{r}{k} (r-k)_j (-n)^{r-k-j} \bar{z}^{r-j-k} f^{(k)}(z) \\ &= (-1)^{r+j} n^j (r)_j \sum_{k=0}^{r-j} \binom{r-j}{k} (-n\bar{z})^{r-j-k} f^{(k)}(z) = n^j (r)_j T_n^{r-j} f. \end{aligned}$$

We integrate by parts in the variable  $z_1$ . Suppose that  $i_1 \leq i_k$ . Using (3.3) and (3.4), we obtain

$$\begin{aligned} &\int_{\mathbb{C}^k} F(z_1, \dots, z_k) [T_n]_{z_1}^{i_1} \overline{[T_n]_{z_2}^{i_1}} K_n(z_1, z_2) \cdots [T_n]_{z_k}^{i_k} \overline{[T_n]_{z_1}^{i_k}} K_n(z_k, z_1) d\mu^k \\ &= \sum_{j=0}^{i_1} \binom{i_1}{j} n^j (i_k)_j \bar{\partial}_{z_1}^{i_1-j} F(z_1, \dots, z_k) \overline{[T_n]_{z_2}^{i_1}} K_n(z_1, z_2) \cdots [T_n]_{z_k}^{i_k} \overline{[T_n]_{z_1}^{i_k-j}} K_n(z_k, z_1) d\mu^k \\ &= \sum_{j=0}^{i_1} \binom{i_1}{j} n^j (i_k)_j \bar{\partial}_{z_1}^{i_k-j} \partial_{z_1}^{i_1-j} F(z_1, \dots, z_k) \overline{[T_n]_{z_2}^{i_1}} K_n(z_1, z_2) \cdots [T_n]_{z_k}^{i_k} K_n(z_k, z_1) d\mu^k. \end{aligned}$$

The last equality holds because  $\overline{[T_n]_{z_2}^{i_1}} K_n(z_1, z_2)$  is analytic in  $z_1$ . If  $i_1 \geq i_k$  instead, the operator acting on  $F$  is  $\sum_{j=0}^{i_k} \binom{i_k}{j} n^j (i_1)_j \bar{\partial}_{z_1}^{i_k-j} \partial_{z_1}^{i_1-j}$ . In any case, the operator acting on  $F$  is  $[\mathcal{D}_{i_k, i_1, n}]_{z_1}$ .

The claim now follows by performing the same procedure in the other variables.  $\square$

Taking  $F = G_k$ , it follows directly from this proposition and (2.4) that

(3.5)

$$(3.6) \quad \begin{aligned} & C_k(\text{trace}_{n,q}g) \\ &= \sum_{0 \leq i_1, \dots, i_k \leq q-1} \frac{n^{-i_1-\dots-i_k}}{i_1! \cdots i_k!} \int_{\mathbb{C}^k} \mathcal{D}_{i_1, \dots, i_k, n} G_k(z_1, \dots, z_k) K_n(z_1, z_2) \cdots K_n(z_k, z_1) d\mu^k(z_1, \dots, z_k). \end{aligned}$$

In the pure polyanalytic case, we have the following, slightly simpler version. Using (3.2), we obtain

(3.7)

$$(3.7) \quad \begin{aligned} & C_k(\text{trace}_{\delta; n, q}g) \\ &= \int_{\mathbb{C}^k} \left( L_{q-1}(-n^{-1}\Delta_{z_1}) \cdots L_{q-1}(-n^{-1}\Delta_{z_k}) G_k(z_1, \dots, z_k) \right) K_n(z_1, z_2) \cdots K_n(z_k, z_1) d\mu_n^k(z_1, \dots, z_k). \end{aligned}$$

#### 4. ESTIMATION OF THE CUMULANTS

In the previous section, we saw that the cumulants  $C_k(\text{trace}_{q,n}g)$  and  $C_k(\text{trace}_{\delta; q, n})$  can both be written as a finite series

$$(4.1) \quad \sum_j n^{-j} \int_{\mathbb{C}^k} H_j(z_1, \dots, z_k) K_n(z_1, z_2) K_n(z_2, z_3) \cdots K_n(z_k, z_1) d\mu_n^k(z_1, \dots, z_k).$$

where the functions  $H_j$  are of the form

$$H_j(z_1, \dots, z_k) = \sum_{\alpha} \prod_{m=1}^k f_{j, \alpha, m}(z_m)$$

Here, the sum over  $\alpha$  is finite and the functions  $f_{j, \alpha, m}$  are bounded and continuous and each of them is either compactly supported or identically equal to 1. In addition, we know that for each  $\alpha$  and  $j$ , the former is the case for at least one  $f_{j, \alpha, m}$ .

In this section, we will show that for every  $j \geq 2$ , the corresponding term in the sum (4.1) tends to zero as  $n \rightarrow \infty$ . Furthermore, we will show that this is the case also for the  $j = 1$  term if  $H_j$  vanishes on the diagonal, i.e. if  $H_j(z, \dots, z) = 0$ . These results are then applied in Section 5 to prove Theorem 1.1.

We will start by applying well-known estimates of the kernel Ginibre kernel  $K_n$ . In fact, essentially the same estimates are also valid for kernels which are defined with more general

than Gaussian weights. We have the following off-diagonal decay estimate for  $K_n$  (Theorem 8.1 in [7]):

$$(4.2) \quad |K_n(z, w)|^2 e^{-n(|z|^2 + |w|^2)} \leq C n^2 e^{-c\sqrt{n} \min\{|z-w|, d(z, \partial\mathbb{D})\}}, \quad z \in \mathbb{D}, w \in \mathbb{C}$$

where the constants  $C$  and  $c$  are absolute constants, in particular independent of  $n$ . We will also need (Proposition 3.6 and p. 1541 in [7])

$$(4.3) \quad |K_n(z, w)|^2 e^{-n|z|^2 - n|w|^2} \leq C n^2 e^{-n(Q(z) - \hat{Q}(z)) - n(Q(w) - \hat{Q}(w))},$$

where  $Q(z) = |z|^2$  and

$$\hat{Q}(z) = \begin{cases} |z|^2, & |z| \leq 1 \\ \log |z|^2 + 1, & |z| \geq 1. \end{cases}$$

According to this estimate,  $|K_n(z, w)|^2 e^{-n|z|^2 - n|w|^2}$  decays quickly to zero as  $n \rightarrow \infty$  if either  $z$  or  $w$  is outside the unit disk.

Let

$$\varepsilon_n = \varepsilon_{n,k} = M_k \frac{\log n}{\sqrt{n}},$$

where  $M_k$  is some large constant to be specified later. We split  $\mathbb{C}^k$  into three different regions  $\Lambda_n, \Gamma_n$  and  $\Omega_n$  as follows:

$$\Lambda_n = \{z_1 \in (1 - \varepsilon_n)\mathbb{D}, \max_{2 \leq j \leq k} |z_j - z_1| \leq \varepsilon_n/2\},$$

$$\Gamma_n = \{d(z_1, \partial\mathbb{D}) \leq \varepsilon_n, (z_2, \dots, z_k) \in \mathbb{C}^{k-1}\},$$

$$\Omega_n = \Omega_n^1 \cup \Omega_n^2 = \{z_1 \in (1 - \varepsilon_n)\mathbb{D}, |z_j - z_1| \geq \varepsilon_n/2 \text{ for some } j\} \cup \{z_1 \in \mathbb{C} \setminus \mathbb{D}, d(z_1, \partial\mathbb{D}) \geq \varepsilon_n\}.$$

The following estimate on the domain  $\Omega_n$  is a straightforward application of the kernel estimates (4.3) and (4.2). We will write

$$R_{k,n}(z_1, \dots, z_k) := K_n(z_1, z_2) \cdots K_n(z_k, z_1).$$

**Proposition 4.1.** *For any  $F \in L^\infty(\mathbb{C}^k)$ , we have that*

$$\int_{\Omega_n} F(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) = O(n^{-1}),$$

whenever  $M_k \geq 2(k+1)/c$ , where  $c$  is the constant in (4.2).

*Proof.* Defining  $f(x) = x - \log x - 1$ , we can use (4.3) to write

$$\begin{aligned} & \left| \int_{\Omega_n^2} F(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) \right| \\ & \leq C^{k/2} n^k \|F\|_\infty \int_{|z|^2 \geq (1+\varepsilon_n)^2} e^{-nf(|z|^2)} dA(z) \left[ \int_{\mathbb{C}} e^{-n(Q(z) - \hat{Q}(z))} dA(z) \right]^{k-1} \end{aligned}$$

For  $x \geq (1 + \varepsilon_n)^2$ , we estimate by convexity

$$f(x) \geq f((1 + \varepsilon_n)^2) + f'((1 + \varepsilon_n)^2)(x - (1 + \varepsilon_n)^2) \geq \varepsilon_n^2 + \frac{2\varepsilon_n + \varepsilon_n^2}{(1 + \varepsilon_n)^2}[x - (1 + \varepsilon_n)^2],$$

so

$$\int_{|z|^2 \geq (1 + \varepsilon_n)^2} e^{-nf(|z|^2)} dA(z) \leq O(e^{-n\varepsilon_n^2}) = O(n^{-M_k^2 \log n}).$$

Because

$$\int_{\mathbb{C}} e^{-n(Q(z) - \hat{Q}(z))} dA(z) = O(1),$$

we obtain

$$\int_{\Omega_n^1} F(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) = O(n^{-\kappa})$$

for any desired  $\kappa > 0$ .

Proceeding to the bulk term, we note that at each point  $(z_1, \dots, z_k) \in \Omega_n^1$ , there is some index  $j$  for which  $|z_j - z_{j+1}| \geq \frac{\varepsilon_n}{2k}$ . Choosing  $j_0 \geq 1$  to be the smallest such index, we observe that  $d(z_{j_0}, \partial\mathbb{D}) \geq \varepsilon_n/2$ . Now we use the off-diagonal estimate (4.2), and obtain

$$|K_n(z_{j_0}, z_{j_0+1})| e^{-n(|z_{j_0}|^2 + |z_{j_0+1}|^2)/2} \leq Cn e^{-\frac{1}{4k} cM_k \log n} = Cn^{1 - \frac{1}{4k} cM_k}.$$

We thus get

$$(4.4) \quad |R_{k,n}(z_1, \dots, z_k) e^{-n(|z_1|^2 + \dots + |z_k|^2)}| \leq Cn^{k - \frac{1}{4k} cM_k}.$$

Because

$$\int_{\mathbb{C} \setminus 2\mathbb{D}} e^{-n(|z|^2 - \log|z|^2 - 1)} = O(e^{-an})$$

for some  $a > 0$ , we see by using (4.3) that

$$\begin{aligned} & \int_{\Omega_n^1} F(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) \\ &= \int_{\Omega_n^1 \cap (2\mathbb{D})^k} F(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) + O(e^{-bn}) \end{aligned}$$

for some  $b > 0$ . Applying this with (4.4) gives

$$\int_{\Omega_n^1} F(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) = O\left(n^{k - \frac{cM_k}{4k}}\right) = O(n^{-1}),$$

where the last equality follows by the choice  $M_k \geq 4k(k+1)/c$ . □

The terms  $j \geq 2$  in (4.1) are negligible because of the following proposition.

**Proposition 4.2.** *Assume that  $f_1 \in L^1(\mathbb{C})$ , and  $f_2, \dots, f_k \in L^\infty(\mathbb{C})$ . Then*

$$\left| \int_{\mathbb{C}^k} \prod_{j=1}^n f_j(z_j) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) \right| \leq n \|f_1\|_{L^1(\mathbb{C})} \prod_{j=2}^k \|f_j\|_\infty.$$

*Proof.* We write the integral on the left hand side as

$$\int_{\mathbb{C}^k} \prod_{j=1}^n f_j(z_j) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) = \int_{\mathbb{C}} f(z_1) F(z_1) dA(z_1),$$

where

$$\begin{aligned} F(z_1) &= \int_{\mathbb{C}^{k-1}} \prod_{j=2}^k f_j(z_j) K_n(z_1, z_2) \cdots K_n(z_k, z_1) d\mu_n^{k-1}(z_2, \dots, z_k) e^{-n|z_1|^2} \\ &= P_n[f_2 P_n[\dots P_n[f_k K_{n,z_1}] \dots]](z_1) e^{-n|z_1|^2}. \end{aligned}$$

Introducing the function

$$F(z_1, z) = P_n[f_2 P_n[\dots P_n[f_k K_{n,z_1}] \dots]](z) e^{-n|z|^2},$$

we may write  $F(z_1) = F(z_1, z_1)$ . Now,

$$|F(z_1, z)| \leq \sqrt{K_n(z, z)} \|F(z_1, \cdot)\|_{L^2(d\mu_n)} e^{-n|z|^2} \leq \sqrt{K_n(z, z)} \sqrt{K_n(z_1, z_1)} \left[ \prod_{j=2}^k \|f_j\|_\infty \right] e^{-n|z|^2}$$

where the first inequality is by Cauchy-Schwarz, and the second follows since projections have norm one. It follows that

$$\left| \int_{\mathbb{C}} f_1(z_1) F(z_1) dA(z_1) \right| \leq \prod_{j=2}^k \|f_j\|_\infty \int_{\mathbb{C}} |f_1(z_1)| K_n(z_1, z_1) e^{-n|z_1|^2} dA(z_1),$$

which by the trivial estimate  $K_n(z, z) e^{-n|z|^2} \leq n$  proves the assertion.  $\square$

The following lemma deals with the term corresponding to  $j = 1$  in (4.1).

**Lemma 4.3.** *Let  $F(z_1, \dots, z_k)$  function of the form*

$$F(z_1, \dots, z_k) = \sum_{\alpha=1}^M \prod_{j \leq k} f_{\alpha,j}(z_j), \quad M \geq 1$$

where  $f_{\alpha,j} \in C^2(\mathbb{C}) \cap L^\infty(\mathbb{C})$ . Assume furthermore that

$$F(z, z, \dots, z) = 0, \quad z \in \mathbb{C}.$$

Then

$$n^{-1} \int_{\mathbb{C}^k} F(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) = o(1).$$

*Proof.* We split the integral over  $\mathbb{C}^k$  as

$$\int_{\mathbb{C}^k} F R_{n,k} d\mu_n^k = \int_{\Lambda_n} F R_{n,k} d\mu_n^k + \int_{\Gamma_n} F R_{n,k} d\mu_n^k + O(n^{-1}).$$

Turning to the integral over  $\Gamma_n$ , we may assume without loss of generality that  $F(z_1, \dots, z_k) = f_1(z_1) \cdots f_k(z_k)$  for some bounded continuous functions  $f_j$ , since vanishing of  $F$  on the diagonal will not be used in this case. It follows immediately from Proposition 4.2 that

$$\begin{aligned} \int_{\Gamma_n} F R_{k,n} d\mu_n &= \int_{\mathbb{C}^k} f_1(z_1) 1_{\{\text{dist}(z_1, \partial\mathbb{D}) \leq M_k \log n / \sqrt{n}\}}(z_1) \prod_{j=2}^k f_j(z_j) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) \\ &\leq O(n \|f_1 1_{\{\text{dist}(z, \partial\mathbb{D}) \leq M_k \log n / \sqrt{n}\}}\|_{L^1}) = O(\sqrt{n} \log n). \end{aligned}$$

For the integral over  $\Lambda_n$ , we estimate  $F$  near  $\Lambda_n$  by a Taylor expansion, and obtain

$$\delta_n := \|F|_{\Lambda_n}\|_{\infty} = (k-1) \max_{2 \leq j \leq k} (\|\partial_{z_j} F\|_{L^{\infty}(\Lambda_n)} + \|\bar{\partial}_{z_j} F\|_{L^{\infty}(\Lambda_n)}) \frac{\varepsilon_n}{2} + O(\varepsilon_n^2) = O(\varepsilon_n).$$

The crude estimate  $|R_{k,n}| e^{-n|z_1|^2 - \dots - n|z_k|^2} \leq n^k$  and  $\text{Vol}(\Lambda_n) = O(\varepsilon_n^{2(k-1)})$  then gives

$$\int_{\Lambda_n} |F| R_{k,n} d\mu_n \leq \int_{\Lambda_n} C \delta_n n^k = O(\delta_n n (M_k \log n)^{2k-2}) = O\left(n \frac{(M_k \log n)^{2k-1}}{\sqrt{n}}\right) = o(n),$$

which completes the proof.  $\square$

## 5. PROOF OF THE MAIN THEOREM

*Proof of Theorem 1.1.* We will start by proving (1.2). By (3.7), we know that  $C_k(\text{trace}_{\delta; n, q})$  can be written as a finite series of the form (4.1). By Proposition 4.2, it is enough pick the terms corresponding to the powers  $n^0$  and  $n^{-1}$  from the expression

$$L_{q-1}^0(n^{-1} \Delta_{z_1}) \cdots L_{q-1}^0(n^{-1} \Delta_{z_k}) G_k(z_1, \dots, z_k).$$

We get

(5.1)

$$\begin{aligned} &C_k(\text{trace}_{\delta; n, q}) \\ &= \int_{\mathbb{C}^k} \left( I + n^{-1}(q-1) \sum_{j=1}^k \Delta_{z_j} \right) G_k(z_1, \dots, z_k) R_{k,n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) + O(n^{-1}). \end{aligned}$$

Suppose first that  $k \geq 3$ . The term corresponding to the identity operator  $I$  in this formula is just the cumulant of order  $k$  in the analytic case  $q = 1$ . By the result of Rider and Virág [15], this expression vanishes as  $n \rightarrow \infty$ . For the  $n^{-1}$  term, it was shown in Lemma 3.3 of [6] that  $\sum_{j=1}^k \Delta_{z_j} G_k(z_1, \dots, z_k)$  vanishes on the diagonal  $z_1 = \dots = z_k$ . The claim now follows by Lemma 4.3.

It remains to calculate the second cumulant, which is equal to the variance. We use (5.1) again. Recall that  $G_2(z, w) = \frac{1}{2}(f(z) - f(w))^2$ . By the result of [15], we have

$$(5.2) \quad \int_{\mathbb{C}^k} G_2(z_1, z_2) R_{2,n}(z_1, z_2) d\mu_n^k(z_1, z_2) \rightarrow \|\bar{\partial}g\|_{L^2(\mathbb{D})}^2 + \frac{1}{2}\|g\|_{H^{1/2}}^2.$$

To analyse the terms involving the Laplacians in (5.1), we calculate

$$(\Delta_z + \Delta_w)(g(z) - g(w))^2 = 2(\Delta g(z) - \Delta g(w))(g(z) - g(w)) + 2|\bar{\partial}g(z)|^2 + 2|\bar{\partial}g(w)|^2.$$

We now observe that  $2(\Delta g(z) - \Delta g(w))(g(z) - g(w))$  vanishes when  $z = w$ . With Lemma 4.3 we now have

$$\begin{aligned} & C_k(\text{trace}_{\delta,n,q}) \\ &= \|\bar{\partial}g\|_{L^2(\mathbb{D})}^2 + \frac{1}{2}\|g\|_{H^{1/2}}^2 + \frac{1}{2n}(q-1) \int_{\mathbb{C}^2} (2|\bar{\partial}g(z)|^2 + 2|\bar{\partial}g(w)|^2 |K_n(z, w)|^2) d\mu^2(z_1, z_2) + o(1) \\ &= \|\bar{\partial}g\|_{L^2(\mathbb{D})}^2 + \frac{1}{2}\|g\|_{H^{1/2}}^2 + \frac{2(q-1)}{n} \int_{\mathbb{C}^2} |\bar{\partial}g(z)|^2 |K_n(z_1, z_2)|^2 d\mu_n^2(z_1, z_2) + o(1) \\ &= \|\bar{\partial}g\|_{L^2(\mathbb{D})}^2 + \frac{1}{2}\|g\|_{H^{1/2}}^2 + \frac{2(q-1)}{n} \int_{\mathbb{C}} |\bar{\partial}g(z)|^2 K_n(z, z) d\mu_n(z) + o(1) \\ &\rightarrow (2q-1) \int_{\mathbb{D}} |\bar{\partial}g|^2 dA + \frac{1}{2}\|g\|_{H^{1/2}}^2. \end{aligned}$$

To obtain the last limit, we used that  $\frac{1}{n}K_n(z, z)d\mu_n(z) \rightarrow 1_{\mathbb{D}}$  weakly.

We will now turn to the proof of (1.3). We have

$$\begin{aligned} & C_k(\text{trace}_{q,n}) \\ &= \sum_{1 \leq i_1, \dots, i_k \leq q-1} \frac{n^{-i_1-\dots-i_k}}{i_1! \dots i_k!} \int_{\mathbb{C}^k} G_k(z_1, \dots, z_k) [T_{n,1}]_{z_1}^{i_1} \overline{[T_{n,1}]_{z_2}^{i_1}} K_n(z_1, z_2) \dots [T_{n,1}]_{z_k}^{i_k} \overline{[T_{n,1}]_{z_1}^{i_k}} K_n(z_k, z_1) d\mu^k \end{aligned}$$

As with pure polyanalytic kernels, we assume first that  $k \geq 3$ . Let us now fix a multi-index  $(i_1, \dots, i_k)$  and estimate the corresponding term in the previous expression. By Proposition 3.1, we get

$$\begin{aligned} & \frac{n^{-i_1-\dots-i_k}}{i_1! \dots i_k!} \int_{\mathbb{C}^k} G_k(z_1, \dots, z_k) [T_{n,1}]_{z_1}^{i_1} \overline{[T_{n,1}]_{z_2}^{i_1}} K_n(z_1, z_2) \dots [T_{n,1}]_{z_k}^{i_k} \overline{[T_{n,1}]_{z_1}^{i_k}} K_n(z_k, z_1) d\mu^k \\ &= \frac{n^{-i_1-\dots-i_k}}{i_1! \dots i_k!} \int_{\mathbb{C}^k} \mathcal{D}_{i_1, \dots, i_k, n} G_k(z_1, \dots, z_k) K_n(z_1, z_2) \dots K_n(z_k, z_1) d\mu^k(z_1, \dots, z_k) \end{aligned}$$

We have

$$n^{-i_1-\dots-i_k} \mathcal{D}_{i_1, \dots, i_k, n} G_k(z_1, \dots, z_k) = O(n^{I(i_1, \dots, i_k)}),$$

where

$$I(i_1, \dots, i_k) := -i_1 - \dots - i_k + \min(i_1, i_k) + \min(i_1, i_2) + \dots + \min(i_{k-1}, i_k).$$

Clearly,  $I(i_1, \dots, i_k) \leq 0$ . We have that  $I(i_1, \dots, i_k) = 0$  if and only if  $i_1 = \dots = i_k$ , and this case is treated above. The terms where  $I(i_1, \dots, i_k) \leq -2$  are negligible by proposition 4.2, so



the only case that needs further analysis is  $I(i_1, \dots, i_k) = -1$ . This happens exactly when there exists a unique  $l \in \{1, 2, \dots, k\}$  such that  $i_l < i_{l+1}$  and for this  $l$  it holds  $i_l = i_{l+1} - 1$ . Here and later, we identify  $i_{k+j}$  with  $i_j$ . This condition on the multi-index  $(i_1, \dots, i_k)$  can be rephrased as follows:  $(i_1, \dots, i_k) = \mathbf{d}(j, r, i)$  for some  $1 \leq j \leq k$ ,  $0 \leq r \leq k-2$  and  $1 \leq i \leq q-1$ , where  $\mathbf{d}(j, r, i)$  is defined to be the multi-index  $(i_1, \dots, i_k)$  which satisfies  $i_m = i$  for  $m \in \{j, \dots, j+r\}$  and  $i_m = i-1$  for  $i_m \in \{i_1, \dots, i_k\} \setminus \{i_j, \dots, i_{j+r}\}$ . If  $(i_1, \dots, i_k) = \mathbf{d}(j, r, i)$ , we have

$$\begin{aligned}
(5.3) \quad & \frac{n^{-i_1-\dots-i_k}}{i_1! \dots i_k!} \mathcal{D}_{i_1, \dots, i_k, n} G_k(z_1, \dots, z_k) \\
&= \frac{n^{-1} i_1! \dots i_{j+r}! (i_{j+r+1} + 1)! \dots i_k!}{i_1! \dots i_k} \partial_{z_j} \bar{\partial}_{z_{j+r+1}} G_k(z_1, \dots, z_k) + O(n^{-2}) \\
&= n^{-1} i \partial_{z_j} \bar{\partial}_{z_{j+r+1}} G_k(z_1, \dots, z_k) + O(n^{-2}),
\end{aligned}$$

where we collected from each operator  $\mathcal{D}_{i_j, i_{j+1}, n}$  the term which has the highest degree in  $n$ . The  $O(n^{-2})$  are negligible by Proposition 4.2. Summing up the contributions from all such  $(i_1, \dots, i_k)$ , we can write the  $k$ 'th cumulant as

$$\begin{aligned}
& \sum_{j=1}^q \int_{\mathbb{C}^k} G_k(z_1, \dots, z_k) K_{\delta; j, n}(z_1, z_2) \dots K_{\delta; j, n}(z_k, z_1) d\mu_n^k(z_1, \dots, z_k) \\
&+ \int_{\mathbb{C}^k} \left[ \sum_{j=1}^k \sum_{r=0}^{k-2} \sum_{i=1}^{q-1} i \partial_j \bar{\partial}_{j+r+1} \right] G_k(z_1, \dots, z_k) R_{k, n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) + O(n^{-1}) \\
&= \sum_{j=1}^q \int_{\mathbb{C}^k} G_k(z_1, \dots, z_k) K_{\delta; j, n}(z_1, z_2) \dots K_{\delta; j, n}(z_k, z_1) d\mu_n^k(z_1, \dots, z_k) \\
&+ \frac{(q-1)(q-2)}{2} \int_{\mathbb{C}^k} \sum_{j \neq l} \partial_j \bar{\partial}_l G_k(z_1, \dots, z_k) R_{k, n}(z_1, \dots, z_k) d\mu_n^k(z_1, \dots, z_k) + O(n^{-1})
\end{aligned}$$

By Lemma 3.4 in [6],  $\sum_{l \neq j} \partial_j \bar{\partial}_l G_k(z_1, \dots, z_k)$  vanishes on  $z_1 = \dots = z_k$ . This together with lemma 4.3 and the result in pure polyanalytic case proves that the cumulants of orders  $k \geq 3$  vanish in the limit.

Next, we calculate the limiting variance. Suppose  $0 \leq i_1 \leq i_2 \leq q-1$ . Again, we only need to analyse the case  $I(i_1, i_2) = -1$ , and this happens exactly when  $i_2 = i_1 + 1$ . Now, using (5.3)

again,

$$\begin{aligned}
& \frac{n^{-2i_1-1}}{i_1!(i_1+1)!} \int_{\mathbb{C}^2} G_2(z_1, z_2) [T_n]_{z_1}^{i_1} \overline{[T_n]_{z_2}^{i_1}} K(z_1, z_2) [T_n]_{z_2}^{i_1+1} \overline{[T_n]_{z_1}^{i_1+1}} K(z_2, z_1) d\mu^2(z_1, z_2) \\
&= n^{-1}(i_1+1) \int_{\mathbb{C}^2} \bar{\partial}_{z_1} \partial_{z_2} G_2(z_1, z_2) K_n(z_1, z_2) K_n(z_2, z_1) d\mu^2(z_1, z_2) + O(n^{-1}) \\
&= -n^{-1}(i_1+1) \int_{\mathbb{C}^2} \bar{\partial}_{z_1} g(z_1) \partial_{z_2} g(z_2) |K_n(z_1, z_2)|^2 d\mu^2(z_1, z_2) + O(n^{-1}) \\
&= -n^{-1}(i_1+1) \int_{\mathbb{C}^2} \left[ \bar{\partial}_{z_1} g(z_1) (\partial_{z_2} g(z_2) - \partial_{z_1} g(z_1)) + |\bar{\partial}_{z_1} g(z_1)|^2 \right] |K_n(z_1, z_2)|^2 d\mu^2(z_1, z_2) + O(n^{-1}) \\
&= -(i_1+1) \int_{\mathbb{D}} |\bar{\partial} g(z)|^2 K_n(z, z) d\mu_n(z) + o(1) \rightarrow -(i_1+1) \int_{\mathbb{D}} |\bar{\partial} g(z)|^2 dA(z)
\end{aligned}$$

Here, the second last equality followed from Lemma 4.3 and integrating out the second variable. For the last limit, we used again the weak convergence  $K_n(z, z) d\mu_n(z) \rightarrow 1_{\mathbb{D}}$ . Summing up all such contributions gives

$$\begin{aligned}
& \int_{\mathbb{C}^2} G_2(z_1, z_2) |K_{q,n}|^2 d\mu^2(z_1, z_2) \\
&= \int_{\mathbb{C}^2} G_2(z_1, z_2) \left[ \sum_{j=1}^q |K_{\delta;j,n}|^2 + 2 \operatorname{Re} \sum_{1 \leq j < l \leq k} K_{\delta;j,n}(z_1, z_2) K_{\delta;l,n}(z_2, z_1) \right] d\mu(z_1, z_2) \\
&\rightarrow \sum_{j=1}^q (2j-1) \|\bar{\partial} g\|_{L^2 D}^2 + \frac{q}{2} \|g\|_{H^{1/2}(\partial \mathbb{D})}^2 - 2 \sum_{j=0}^{q-2} (j+1) \|\bar{\partial} g\|^2 = q(\|\bar{\partial} g\|_{L^2(\mathbb{D})}^2 + \frac{1}{2} \|g\|_{H^{1/2}(\partial \mathbb{D})}^2)
\end{aligned}$$

□

## REFERENCES

- [1] LD Abreu, P Balazs, M De Gosson, and Z Mouayn, *Discrete coherent states for higher landau levels*, Annals of Physics **363** (2015), 337–353.
- [2] Luís Daniel Abreu, *Sampling and interpolation in bargmann–fock spaces of polyanalytic functions*, Applied and Computational Harmonic Analysis **29** (2010), no. 3, 287–302.
- [3] Luis Daniel Abreu and Hans G Feichtinger, *Function spaces of polyanalytic functions*, Harmonic and complex analysis and its applications, Springer, 2014, pp. 1–38.
- [4] Luís Daniel Abreu, João M Pereira, José Luis Romero, and Salvatore Torquato, *The weyl-heisenberg ensemble: hyperuniformity and higher landau levels*, arXiv preprint arXiv:1612.00359 (2016).
- [5] Yacin Ameur, Haakan Hedenmalm, Nikolai Makarov, et al., *Random normal matrices and ward identities*, The Annals of Probability **43** (2015), no. 3, 1157–1201.
- [6] Yacin Ameur, Håkan Hedenmalm, Nikolai Makarov, et al., *Fluctuations of eigenvalues of random normal matrices*, Duke mathematical journal **159** (2011), no. 1, 31–81.
- [7] Yacin Ameur, Nikolai Makarov, and Håkan Hedenmalm, *Berezin transform in polynomial bergman spaces*, Communications on Pure and Applied Mathematics **63** (2010), no. 12, 1533–1584.
- [8] Mark Benevich Balk and MF Zuev, *On polyanalytic functions*, Russian Mathematical Surveys **25** (1970), no. 5, 201.

- [9] Ovidiu Costin and Joel L Lebowitz, *Gaussian fluctuation in random matrices*, Physical Review Letters **75** (1995), no. 1, 69.
- [10] Nour eddine Askour, Ahmed Intissar, and Zouhair Mouayn, *Espaces de bargmann généralisés et formules explicites pour leurs noyaux reproduisants*, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics **325** (1997), no. 7, 707–712.
- [11] Antti Haimi, *Bulk asymptotics for polyanalytic correlation kernels*, Journal of Functional Analysis **266** (2014), no. 5, 3083–3133.
- [12] Antti Haimi and Haakan Hedenmalm, *The polyanalytic ginibre ensembles*, Journal of statistical physics **153** (2013), no. 1, 10–47.
- [13] John Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág, *Zeros of gaussian analytic functions and determinantal point processes*, vol. 51, American Mathematical Society Providence, RI, 2009.
- [14] Odile Macchi, *The coincidence approach to stochastic point processes*, Advances in Applied Probability (1975), 83–122.
- [15] Brian Rider and Bálint Virág, *The noise in the circular law and the gaussian free field*, International Mathematics Research Notices **2007** (2007), rnm006.
- [16] Tomoyuki Shirai, *Ginibre-type point processes and their asymptotic behavior*, Journal of the Mathematical Society of Japan **67** (2015), no. 2, 763–787.
- [17] Alexander Soshnikov, *Determinantal random point fields*, Russian Mathematical Surveys **55** (2000), no. 5, 923–975.
- [18] NL Vasilevski, *Poly-fock spaces*, Differential Operators and Related Topics, Springer, 2000, pp. 371–386.

FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA  
 OSKAR-MORGENSTERN-PLATZ 1, A-1090 VIENNA, AUSTRIA

DEPARTMENT OF MATHEMATICS, KTH ROYAL INSTITUTE OF TECHNOLOGY,  
 STOCKHOLM, 100 44, SWEDEN